

Conjugate Directions in Lattice Landau and Coulomb Gauge Fixing

R.J. Hudspith*

*Department of Physics and Astronomy, York University,
Toronto, Ontario, M3J 1P3, Canada
E-mail: renwick.james.hudspith@gmail.com*

We provide details expanding on our implementation of a non-linear conjugate gradient method with Fourier acceleration for lattice Landau and Coulomb gauge fixing. We find clear improvement over the Fourier accelerated steepest descent method, with the average time taken for the algorithm to converge to a fixed, high accuracy, being reduced by a factor of 2 to 4. We show such improvement for the logarithmic definition of the gauge fields here, having already shown this to be the case for a more common definition. We also discuss the implementation of an optimal Fourier accelerated steepest descent method.

*The 32nd International Symposium on Lattice Field Theory,
23-28 June, 2014
Columbia University New York, NY*

*Speaker.

1. Introduction

Fixing to the smooth, covariant, Coulomb and Landau gauges can be useful for lattice field theorists. The matching of lattice results to continuum perturbation theory, of particular interest for vertex functions, demands the use of a fixed gauge, usually Landau. Coulomb gauge has other benefits; it can be used for the creation of gauge-fixed wall source fermion propagators or for measurements of the static quark potential.

The Landau and Coulomb gauges are two of the easiest (that have useful continuum analogs) to implement on the lattice as their gauge conditions can be defined as the minimisation of a functional. Such a description allows us to define a smooth deformation of the fields in which we can systematically move toward a local minimum of the functional. This is the aim of many numerical optimisation techniques and several methods are available [1, 2, 3].

Conjugate gradient methods were first introduced in [4] to solve systems of equations and were generalised to the associated problem of solving non-linear, non-quadratic functions in [5, 6, 7]. We show how a non-linear conjugate gradient method can be used with Fourier Acceleration to create a very fast gauge fixing procedure for both Coulomb and Landau gauge with exact and approximate definitions of the gauge field. We also comment on the ineffectiveness of fixing to these gauges using a so-called “optimal” steepest descent (SD) routine that tunes the step-size to best minimise the functional at each iteration.

2. Fields and Definitions

Lattice gauge fields (the A_μ 's) of an N_d -dimensional, strongly interacting gauge theory are formally defined by their path-ordered parallel transport matrices (known as lattice links),

$$U_\mu \left(x + a \frac{\hat{\mu}}{2} \right) = P \left[e^{ig_0 \int_x^{x+a\hat{\mu}} A_\mu(x) dx_\mu} \right] \approx e^{iag_0 A_\mu \left(x + a \frac{\hat{\mu}}{2} \right)}, \quad (2.1)$$

and are well described as living halfway between the sites x and $x + a\hat{\mu}$ of the lattice, where a is the lattice spacing, g_0 is the bare coupling and μ is the field's polarisation direction. The links are elements of the group $SU(N_C)$, the fields are elements of the Lie Algebra $\mathfrak{su}(N_C)$ with N_C being the number of color charges of the theory.

The fields cannot be transformed in a gauge-covariant manner directly in lattice simulations, instead they have to be altered by gauge transformation,

$$U_\mu^{(g)} \left(x + a \frac{\hat{\mu}}{2} \right) = g(x) U_\mu \left(x + a \frac{\hat{\mu}}{2} \right) g(x + a\hat{\mu})^\dagger. \quad (2.2)$$

Such a transformation preserves the gauge invariance of the action.

The techniques we consider in this work for fixing the gauge (gradient descent methods) require access to the gauge fields directly. There is some freedom in the choice of field definition: a common one is what we call the “Hermitian Projection” and is accurate only up to terms of

$$O\left(A_\mu\left(x+a\frac{\hat{\mu}}{2}\right)^3\right),$$

$$\begin{aligned} ag_0 A_\mu\left(x+a\frac{\hat{\mu}}{2}\right) &= \frac{U_\mu\left(x+a\frac{\hat{\mu}}{2}\right) - U_\mu\left(x+a\frac{\hat{\mu}}{2}\right)^\dagger}{2i} \\ &\quad - \frac{1}{2iN_C} \text{Tr}\left[U_\mu\left(x+a\frac{\hat{\mu}}{2}\right) - U_\mu\left(x+a\frac{\hat{\mu}}{2}\right)^\dagger\right] \cdot I_{N_C \times N_C}. \end{aligned} \quad (2.3)$$

Some authors have also investigated the use of exact logarithmic fields [8, 9]. These will be used here to illustrate the general performance properties of our Fourier Accelerated Conjugate Gradient algorithm.

Cayley-Hamilton theorem states that every matrix is a solution of its own characteristic equation. In terms of our links the following finite expansion can be used (where the f 's are necessarily complex and I is the identity matrix) [10],

$$U_\mu\left(x+a\frac{\hat{\mu}}{2}\right) = f_0 \cdot I_{N_C \times N_C} + f_1 A_\mu\left(x+a\frac{\hat{\mu}}{2}\right) + \cdots + f_{N_C-1} A_\mu\left(x+a\frac{\hat{\mu}}{2}\right)^{N_C-1}. \quad (2.4)$$

The f 's are the solution to the following Vandermonde equation,

$$\begin{pmatrix} 1 & q_0 & q_0^2 & \cdots & q_0^{N_C-1} \\ 1 & q_1 & q_1^2 & \cdots & q_1^{N_C-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & q_{N_C-1} & q_{N_C-1}^2 & \cdots & q_{N_C-1}^{N_C-1} \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_{N_C-1} \end{pmatrix} = \begin{pmatrix} e^{iq_0} \\ e^{iq_1} \\ \vdots \\ e^{iq_{N_C-1}} \end{pmatrix}. \quad (2.5)$$

The q 's are the Eigenvalues of $A_\mu\left(x+a\frac{\hat{\mu}}{2}\right)$ and the e^{iq} 's are the Eigenvalues of the link matrix $U_\mu\left(x+a\frac{\hat{\mu}}{2}\right)$. For our logarithm, we compute the e^{iq} 's from the characteristic equation of the link matrix $U_\mu\left(x+a\frac{\hat{\mu}}{2}\right)$ and take their complex argument to define the q 's on the principal branch. The solution to general Vandermonde systems can become numerically unstable, but for SU(3) well-behaved functions for the f 's in terms of the q 's can be used [10]. This leads to a computationally efficient and numerically stable definition for the principal logarithm of SU(3) matrices,

$$ag_0 A_\mu\left(x+a\frac{\hat{\mu}}{2}\right) = \frac{f_2^* U_\mu\left(x+a\frac{\hat{\mu}}{2}\right) - f_2 U_\mu^\dagger\left(x+a\frac{\hat{\mu}}{2}\right) - 2i\Im(f_0 f_2^*) \cdot I_{3 \times 3}}{2i\Im(f_1 f_2^*)}.$$

We define the lattice finite difference of the gauge fields by,

$$a\Delta_\mu A_\mu(x) = \sum_\mu \left[A_\mu\left(x+a\frac{\hat{\mu}}{2}\right) - A_\mu\left(x-a\frac{\hat{\mu}}{2}\right) \right]. \quad (2.6)$$

It will be useful to monitor our algorithm (at its n^{th} iteration) by defining the quantity (V is the lattice volume),

$$\Theta^{(n)} = \frac{1}{N_C V} \sum_x \text{Tr} \left[\left(a\Delta_\mu ag_0 A_\mu^{(n)}(x) \right)^2 \right]. \quad (2.7)$$

3. Fourier Accelerated Steepest Descent

The continuum gauge condition for Landau gauge (and Coulomb with μ only in the spatial directions) is $\partial_\mu A_\mu(x) = 0$, and the lattice analog of this is $a\Delta_\mu ag_0 A_\mu(x) = 0$. This occurs at the minimum of the gauge fixing functional,

$$F(U) = \frac{1}{N_C N_d V} \sum_{x,\mu} \text{Tr} \left[A_\mu \left(x + a \frac{\hat{\mu}}{2} \right)^2 \right]. \quad (3.1)$$

A general steepest descent step that drives the gauge fields toward a local minimum of the functional (and thus reduces $a\Delta_\mu ag_0 A_\mu(x)$ to 0) is,

$$g(x) = e^{-i\alpha a\Delta_\mu ag_0 A_\mu(x)}, \quad (3.2)$$

followed by a gauge transformation of the links (Eq.2.2). The parameter α is a small, fixed, positive tuning parameter. This procedure is repeated until some convergence criterion is met.

The authors of [11] suggested that a re-scaling of the Eigenvalues of the Abelian Laplacian operator ($\Delta^2 A_\mu(x)$) in momentum space was sufficient to reduce some of the critical slowing down in each steepest descent iteration, this is effectively a preconditioner of the gradient. A Fourier Accelerated Steepest Descent (FASD) iteration is,

$$g(x) = e^{\tilde{F} \frac{-i\alpha \hat{p}_{\text{Max}}^2}{V \hat{p}^2} F a\Delta_\mu ag_0 A_\mu(x)}. \quad (3.3)$$

Followed by a gauge transformation. \hat{p}_{Max}^2 takes the values $4N_d$ and $4(N_d - 1)$ for fixing to Landau and Coulomb gauge respectively. F and \tilde{F} are forward and backward Fast Fourier Transforms (FFTs using FFTW [12] in our implementation). We use the definition of lattice momentum (the n_μ 's are Fourier modes, L_μ is the lattice length in the μ direction),

$$\hat{p}^2 = 2 \left(N_d - \sum_\mu \cos \left(\frac{2\pi n_\mu}{L_\mu} \right) \right). \quad (3.4)$$

The rest of this paper will present results from gauge group $\text{SU}(3)$, $N_d = 4$ dimensional lattice gauge theory. In particular, $N_f = 2 + 1$, Iwasaki gauge, DWF configurations are used. We will consider three different lattice volumes; the $32^3 \times 64$ data has a lattice spacing of $a^{-1} \approx 2.38$ GeV, the $24^3 \times 64$ and $16^3 \times 32$ configurations' $a^{-1} \approx 1.79$ GeV [13].

3.1 Optimal FASD

Many popular techniques for fixing the gauge require some form of tuning to be effective, with FASD and over-relaxation being examples. The tuning parameter can depend on the bare coupling, the gauge action, the field definition and the functional being used. If the tuning parameter is too large the routine will not converge, too small and the routine will have poor performance. The optimal FASD routine, and the Fourier Accelerated Conjugate Gradient algorithm discussed later do not need to be tuned.

We can define a so-called ‘‘optimal’’ steepest descent routine by choosing the parameter α_n to be the one which best minimises the gauge functional (Eq.3.1) at each iteration of the algorithm.

Evaluating the optimal parameter α_n exactly (by e.g. a binary search) would be prohibitively expensive as each evaluation of what we call a probe α' requires both a lattice-wide gauge transformation and subsequent functional evaluation. The approach we take instead is to evaluate a selection of probe α' 's and create a cubic spline interpolation of the probes for which we can solve for the minimum and hence provide a good approximation to the actual minimum. Such a process is called an approximate line search and will be integral to the non-linear conjugate-gradient method discussed in the next section.

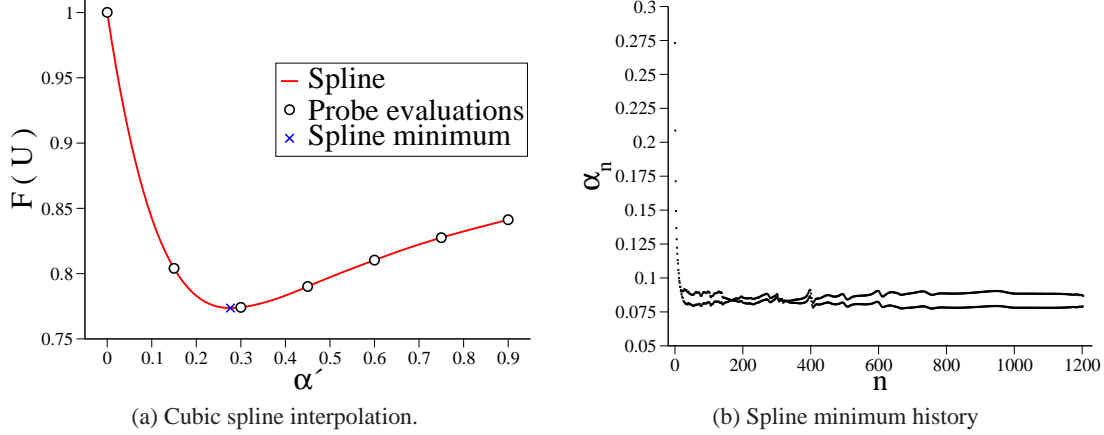


Figure 1: Fig.1a shows a cubic spline interpolation and the computation of the spline minimum for a series of probe α' 's. Fig.1b shows the history of the spline minimum per iteration of the optimal FASD from an example $16^3 \times 32$ configuration converged to a high accuracy of $\Theta^{(n)} < 10^{-14}$.

Fig.1a shows the cubic spline interpolation for a group of probe evaluations for an FASD step. Fig.1b illustrates the typical history of the cubic spline minimum per iteration of the optimal FASD Landau gauge fixing routine, taken from a single $16^3 \times 32$ configuration. We see that after the first 50 or so iterations the best α_n is roughly constant at around 0.08 which is consistent with the best tuning parameter of similar configurations found in the literature [11, 9]. The approximate optimal α_n oscillates each iteration around a central value, illustrating the routine correcting itself from an over or under-estimate of the true optimal α_n .

As the optimal FASD shows a long plateau of the best α_n with little variation the cost of the line search is not worthwhile for the FASD algorithm. This is because it is roughly $N \times$ the cost of one fixed- α iteration (N is the number of probe α' 's) and we need at least $N = 3$ to compute a cubic spline interpolant. If the number of iterations is not reduced by a factor of N the optimal FASD is less optimal (in the sense of time taken to reach a desired precision) than the fixed- α FASD, which we find to be the case.

4. Conjugate Directions in Gauge Fixing

A detailed description of the of the Fourier Accelerated Conjugate Gradient (FACG) algorithm is presented in Alg.1 and Alg.2 of [14] and will not be reproduced here.

Fig.2a shows a histogram of α_n taken from the probes (0.0, 0.15, 0.3, 0.45, 0.6) for every iteration of the configuration used in the previous section. We can see that most of the optimal α_n 's lie

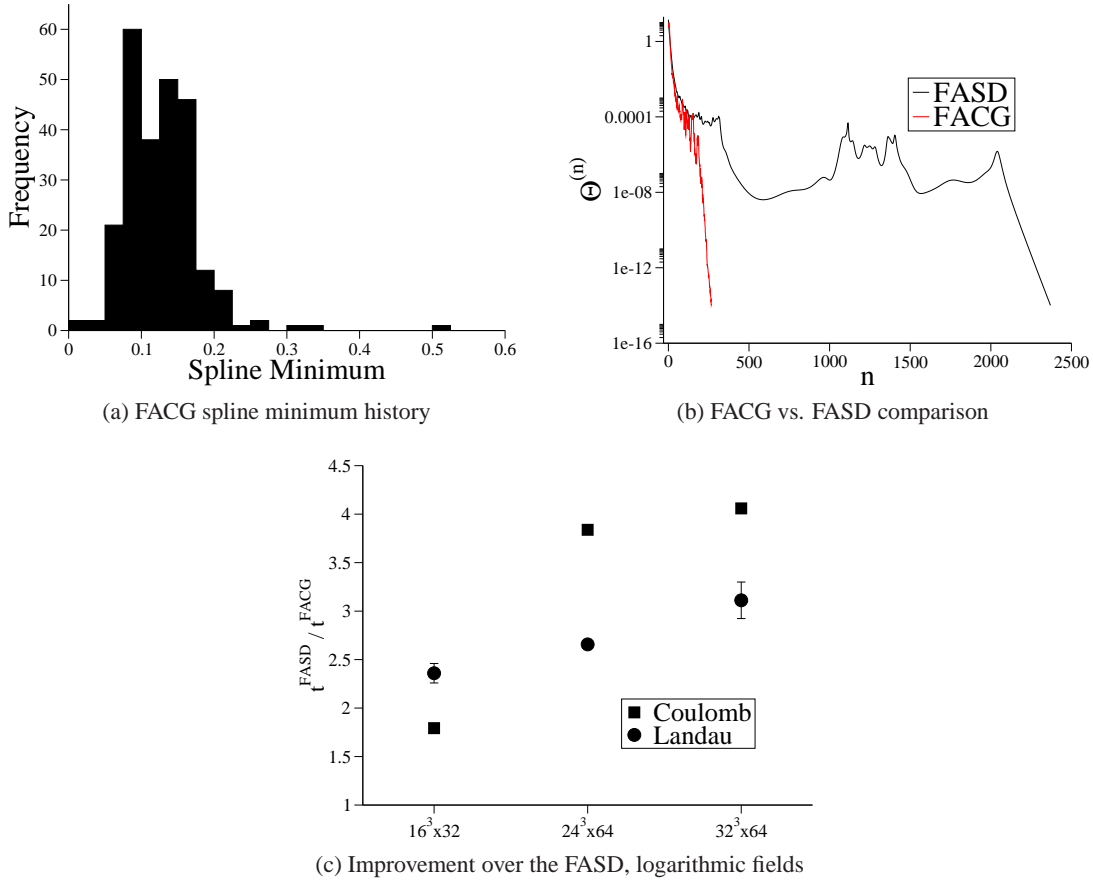


Figure 2: Fig.2a shows the optimal α_n 's of the FACS algorithm for an example $16^3 \times 32$ configuration being fixed to $\Theta^{(n)} < 10^{-14}$. Fig.2b shows the number of iterations required for the (fixed- α) FASD and FACS algorithms to reach the same accuracy for the same configuration, using the Hermitian Projection definition of the gauge fields. Fig.2c illustrates a measure of the effective speed up of our FACS compared to our FASD routine for the logarithmic field definition for 3 different lattice volumes.

in the range $[0, 0.3]$ and so we choose the three probe evaluations $(0.0, 0.15, 0.3)$ for our routine. The probe at 0.0 is very beneficial in bounding the minimum and is cheap to evaluate as no gauge transformations are needed. If the spline minimum is negative, we set α_n to 0.0 (i.e. we generate a new descent direction) or if it is greater than 0.3 we set it to 0.3 knowing that the algorithm will accommodate for a sub-optimal step.

We provide a comparison of the FACS algorithm (using the 3 probes) to the fixed- α FASD algorithm in Fig.2b showing the number of iterations required to reach $\Theta^{(n)} < 10^{-14}$ for a single configuration, although this behaviour is ubiquitous. We see that for this configuration there is a factor of around 10 reduction in the number of iterations to reach a high level of convergence. As mentioned for the optimal FASD routine, if the factor of reduction in the number of iterations is greater than the number of probe evaluations (N) then the routine is beneficial.

In [14](Fig.1) we show for the Hermitian Projection definition of the fields a factor of 2 to 4 improvement in the average time taken to converge to a high fixed accuracy was obtained for the FACS routine compared to the FASD. In Fig.2c, we show that such speedup is also available for

the logarithmic field definition for the three volumes considered. The measure of improvement was taken to be the time taken for 25 randomly gauge transformed copies of the same configuration at each volume for the FASD algorithm divided by the time taken for the FACG algorithm.

5. Conclusions

We have presented results using the FACG algorithm for fixing to Landau and Coulomb gauge. It is shown to be faster than our implementation of the FASD algorithm for both the logarithmic and Hermitian Projection field definitions. We have illustrated why an optimal FASD algorithm does not provide a faster method than the fixed- α procedure, and have shown why the literature value of $\alpha = 0.08$ is often a good choice for the tuning parameter, although the exact optimal tuning parameter is dependent on several factors. Unlike the fixed- α FASD or over-relaxation methods, the FACG does not require one to tune a parameter.

The FACG algorithm converges in fewer iterations than the FASD, and requires only one set of FFTs per iteration; if FFTs are particularly costly this method will show greater improvement over the FASD than presented here. The FACG method (as it is a gradient descent method) allows for options where standard relaxation techniques cannot be used; different gauge field definitions or alternate functionals can be accommodated easily.

6. Acknowledgements

The results produced in this work were generated on Columbia University's cluster "CUTH" using the package GLU (<https://github.com/RJHudspith/GLU>). RJH would like to thank R. Lewis. RJH is supported by the Natural Sciences and Engineering Research Council of Canada.

References

- [1] J. Mandula and M. Ogilvie *Phys.Lett.* **B185** (1987) 127–132.
- [2] J. E. Mandula and M. Ogilvie *Phys.Lett.* **B248** (1990) 156–158.
- [3] P. de Forcrand *Nucl.Phys.Proc.Suppl.* **9** (1989) 516–520.
- [4] M. R. Hestenes and E. Stiefel *J. Res. Natl. Bur. Stand.* **49** (1952) 409–436.
- [5] B. Polyak {USSR} *Computational Mathematics and Mathematical Physics* **9** (1969), no. 4 94–112.
- [6] R. Fletcher and C. M. Reeves *The Computer Journal* **7** (1964), no. 2 149–154.
- [7] R. G. Polak, E. *ESAIM: Mathematical Modelling and Numerical Analysis* **3** (1969), no. R1 35–43.
- [8] S. Furui and H. Nakajima *Nucl.Phys.Proc.Suppl.* **73** (1999) 865–867.
- [9] E.-M. Ilgenfritz et al. *Phys.Rev.* **D83** (2011) 054506.
- [10] C. Morningstar and M. Peardon *Phys. Rev. D* **69** (Mar, 2004) 054501.
- [11] C. T. H. Davies et al. *Phys. Rev. D* **37** (Mar, 1988) 1581–1588.
- [12] M. Frigo and S. G. Johnson *Proceedings of the IEEE* **93** (2005), no. 2 216–231.
- [13] **UKQCD collaborations** Collaboration, RBC et al. [arXiv:1411.7017](https://arxiv.org/abs/1411.7017).
- [14] R. Hudspith *Computer Physics Communications* **187** (2015), no. 0 115–119.